基于D-S证据推理方法实现异构动态安全信息的融合

徐欢潇[1]　吉　雅[2]　陈虹云[1]　祁　娟[3]

1. 南通理工学院 计算机与信息工程学院　江苏南通　226019

2. 南通理工学院 信息化建设与管理处　江苏南通　226019

3. 南通理工学院 传媒与设计学院　江苏南通　226019

摘　要：随着网络规模的不断扩大，网络入侵手段越来越多样化和智能化，基于网络入侵检测、网络防火墙、防病毒系统和终端监控系统等一系列安全设备的纵深防御系统已经建立起来，在此过程中暴露出一些问题和研究趋势。采用扩展的D-S证据推理方法实现了异构动态安全信息的融合。

关键词：多源异构；数据融合；移动网络；钓鱼检测

Implementation of Heterogeneous Dynamic Security Information Fusion Based on D-S Evidence Reasoning Method

Huanxiao Xu [1], Ya Ji[2], Hongyun Chen[1], Juan Qi[3]

1. School of Computer and information engineering, Nantong Institute of Technology, Jiangsu 226019, China

2. Information Construction and Management Department of Nantong Institute of Technology, Nantong 226019, Jiangsu, China

3. School of Media and Design, Nantong Institute of Technology, Nantong 226019, Jiangsu, China

Abstract: With the continuous expansion of network scale, network intrusion methods are becoming increasingly diverse and intelligent. A series of security devices such as network intrusion detection, network firewalls, antivirus systems, and terminal monitoring systems have been established for in-depth defense systems, which have exposed some problems and research trends in this process. The fusion of heterogeneous dynamic security information was achieved using an extended D-S evidence reasoning method.

Keywords: Multi source isomerism; Data fusion; Mobile network; Fishing detection

依托项目:2022年度江苏省高等学校基础科学（自然科学）研究面上项目“基于多源异构数据融合的移动终端网络钓鱼检测研究”（编号：22KJD520009）

1　引言

随着计算机网络的不断扩展，对网络安全的要求也不断提高。一方面，入侵检测系统存在误报率高、报警重叠、漏报、报警语义弱等问题，给网络入侵的及时识别、分析和响应带来很大困难和误导性影响。另一方面，随着安全系统及设备的增加，各种报警信息和日志数量也在成倍增加。原始报警信息基本上是底层信息，过于简单和冗余，还存在虚警问题。因此，迫切需要有效的异构安全信息分析与处理技术。

2　网络数据安全防护机制总体结构

本文基于网络安全数据服务可用性和性能的可视化设计了安全保护机制，并采用客户机/服务器模式设计了系统的

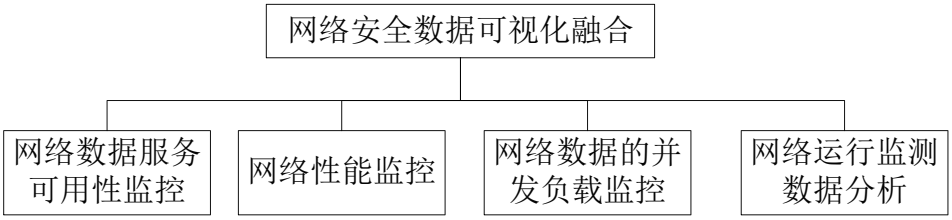

图1　网络数据安全防护机制总体结构

软件部分。图1显示了网络数据安全保护机制的结构[1]。

本文设计的网络数据安全防护机制的功能是：

（1） 有效保障网络流量数据的安全，统计网络访问者的浏览情况，管理用户下载文件、页面显示、网络文件等的次数；

（2） 确保网络用户操作过程中安全数据的稳定性，实现对用户在线类型、操作系统和浏览器版本的分析，确保系统的稳定性；

（3）管理用户来源，确保网络用户在运营过程中的安全，全面监控网络用户的运营[2]；

（4）维护网络数据传输的安全功能。检测单个IP和IP段网络主机的生存能力，检测结果可以反映主机的生存能力。

3 模糊D-S证据论的基础理论[3]

所谓证据论是用于描述或量化证据的一组函数理论，这些函数是概率分配(BPA)函数、信任(Bel)函数以及似真函数(Pls)等等。

一条U上的模糊D-S证据体B具有下列形式：

B={(f1，m1),(f2，m2)，…，(fa ，ma)}其中fi∈U，称为B的模糊焦点元；mi为相应焦点元的密度，也称为概率分配(BPA)函数。

定义1 ：U上的概率分配(BPA)函数 m为一映射m：U→[0，1]，满足 m(f)≥0，且$\sum_{f\in U} m(f)=1$,m(f)=1，如果 m(f)>0，则称 f 为模糊焦点元。

定义2 ：一个函数 m(x)是识别框U幂集2^U上的一个映射：2^U→[0，1]是一个基本的概率分配(BPA)函数，如果它满足：m(ϕ)=0；$\sum_{x\subseteq U} m(x)=1$。

定义 3：信任函数(Bel)：2^U→[0，1]，满足：Bel(A)=$\sum_{x\subseteq A} m(x)$，对于所有 A⊆U 是一个信任函数，为A的可信度，即事件A确认出现的最小概率。

定义 4:设m为U上的概率分配(BPA)函数，则U上的相应于m的似然函数 Pls为一映射 Pls：U→[0，1]，满足：Pls(A)=$\sum_{x\subseteq U}\sum_{x\cap A\neq\phi} m(x)$对所有 A⊆U，为$\overline{A}$不可能断定的最大概率。

定义 5 ：Dempster- Shafter 的组合规则如下：

设m_1和m_2是2^U上的概率分配函数，假设 F=$\sum_{x\cap y=\phi} m_1(x)m_2(y)<1$，记 N=$\sum_{x\cap y=\phi} m_1(x)m_2(y)$，则 m($\phi$)=0，且对于U的所有子集$A\neq\phi$ ，m(A)=$(1/N)\sum_{x\cap y=A} m_1(x)m_2(y)$，若m是 2^U→[0，1]的一个概率分配函数，则m称 m_1和 m2的正交和，记$m_1\oplus m_2$，如果 N=$\sum_{x\cap y=\phi} m_1(x)m_2(y)=0$则正交和 $m_1\oplus m_2$不存在，且m_1和m_2是完全冲突的，K=1/N 称m_1和m_2的正交和规范常数，是测量两个概率分配函数之间的冲突范围，在融合源 m>2 时 Dempster- Shafter组合规范的推广，遵守$m=m_1\oplus m_2...\oplus m_n=((((m_1\oplus m_2)\oplus m_3)...)\oplus m_n)$的融合规则，计算组合的BPAm。

4 多源数据融合

多源融合模块将融合多个决策引擎的输出结果，从全局的角度分析网络状况，

进一步提升识别网络攻击类型的效果。然而当各决策引擎输出结果（多源融合中称作证据）之间存在冲突时，会严重影响融合效果，因此首先要对冲突证据进行修正，再进行多源数据融合。

多源数据融合将融合多个决策引擎的输出结果，综合分析网络安全状况。目前常用于网络安全态势评估中的融合算法主要是 D-S 证据理论。D-S证据理论是一个经典的融合算法，通过对相同的事件进行关联、融合，得出一致的结论：

$$m(A)=\frac{\sum_{B_i\cap C_j=A} m_1(B_i)m_2(C_j)}{1-\sum_{B_i\cap C_j=\phi} m_1(B_i)m_2(C_j)} \quad (1)$$

其中$\sum_{B_i\cap C_j=A} m_1(B_i)m_2(C_j)$称为冲突系数，$B_i$和$C_j$为焦元，m(A)表示融合后的结果。

如表1所示[4]，某次实验中，决策引擎1输出结果为：75%的概率为Dos攻击，15%的概率为Backdoor攻击，10%的概率为Shellcode攻击。决策引擎2输出结果为：80%的概率为Dos攻击，20%的概率为Backdoor攻击。使用D-S证据理论融合决策引擎1和决策引擎2的输出结果：95%的概率为Dos攻击，5%的概率为Backdoor攻击。

表 1 D-S 证据融合

	Dos	Backdoor	Shellcode
决策引擎1	0.75	0.15	0.1
决策引擎2	0.8	0.2	0
D-S融合	0.95	0.05	0

多证据融合的目标是融合多个决策引擎输出的证据，然而当它们之间存在冲突时，会严重影响融合效果。因此多证据融合首先进行冲突证据修正，减小证据之间的矛盾性，再使用基于指数加权的D-S充分融合修正后的证据。

D-S证据理论在处理数据融合中有着很多优势：可以通过积累证据不断的缩小假设的范围；对因为随机和模糊产

生的不确定性有较好的处理能力；可以不需要条件概率和先验概率的密度。其中最大优势就在于解决了不确定性计算的问题。

不同的传感器对威胁精度也各不相同，因此需要对D-S证据理论给出的概率予以不同的加权。

5 基于异构动态安全信息的报警融合

D-S证据理论是Dempster于1967年首先提出，由他的学生Shafer于1976年进一步发展起来的一种不精确推理理论，也称为Dempster/Shafer 证据理论(D-S证据理论)[5]。

D-S包含规则同样信任不同传感器提供的证据。然而，实际情况并非如此。我们观察到，对于同类传感器，本地传感器应该比远程传感器提供更可靠的信息；即使同一传感器安装在网络的不同位置，它也将具有不同的检测能力；不同类型的传感器对同一种犁的攻击具有不同的检测能力和准确性，为融合判断提供的证据的重要性和可靠性也有很大差异。

为了解决上述问题，本文利用扩展的证据组合规则对不同传感器的证据赋予不同的权重，即对不同传感器赋予不同的信任。文献[6]建议在内容感知研究领域通过比例加权来改进证据组合规则。

组合规则如下：

$$m_{12}(A)=\frac{\sum_{A_i\cap A_j=A}\left(w_1m_1(A_i)\cdot w_2m_2(A_j)\right)}{\sum_{A_i\cap A_j\neq\varnothing}\left(w_1m_1(A_i)\cdot w_2m_2(A_j)\right)} \tag{2}$$

可以证明，组合概率分布函数不满足 $\sum_{A\subseteq2}m(A)=1$ ； 它不能用于证据积累。本文采用指数加权，如公式(3)所示。通过采用指数加权规则（证明过程省略），可以证明组合证据仍然满足概率分布函数的基本性质。

$$m_{12}(A)=\frac{\sum_{A_i\cap A_j=A}m_1[(A_i)]^{w_1}m_2[(A_j)]^{w_2}}{\sum_{A_i\cap A_j\neq\varnothing}m_1[(A_i)]^{w_1}m_2[(A_j)]^{w_2}} \tag{3}$$

w_2 是观察到的 O_i 的重量， 而且当 $w_1=w_2$， 公式(3)简化为基本的D-S组合规则。权重可以通过基于最大熵准则或最小均方误差准则的训练样本来估计。

6 总结

在网络发展过程中，网络安全是一个特别重要的限制因素。本文提出了一种扩展的D-S证据推理方法，实现了异构动态安全信息的融合。

参考文献：

[1]Huanxiao Xu et al. Research on Mobile Phishing Detection Association Technology Based on Multi-Source Heterogeneous Security Data Fusion[C]. Huhhot, China, 2021.

[2]Ovando D, Poon S, Costello C. Opportunities and precautions for integrating cooperation and individual transferable quotas with territorial use rights in fisheries. Bulletin of Marine Science -Miami-, vol. 93, no. 1, pp. 101-115, 2017.

[3]Dempster A.Upper and lower probabilities induced by multivalued mapping[J].Annals of Mathematical Statistics,1967,38(2):325- 339.

[4]田晓雄.多源异构数据融合的网络安全态势评估研究[D].山西财经大学,2021.DOI:10.27283/d.cnki.gsxcc.2021.000279.

[5]Wang W. Government ecological governance management based on heterogeneous multi-processors and dynamic image sampling - ScienceDirect. Microprocessors and Microsystems, vol. 82, no. 1, pp. 103909, 2021.

作者简介：

徐欢潇(1991.3-) ，女，汉族，江苏南通人，南通理工学院，讲师，硕士学位，专业：信息与通信工程，研究方向：网络安全。

吉雅（1998.8-），女，汉族，江苏盐城人，南通理工学院，本科学位，专业：软件工程，研究方向：软件工程方法与技术。

陈虹云（1992.5-），女，汉族，江苏南通人，南通理工学院，讲师，硕士学位，专业：信息与通信工程，研究方向：信息安全。

祁娟（1993.10-），女，汉族，江苏淮安人，南通理工学院，讲师，硕士学位，专业：数字媒体艺术，研究方向：数字媒体艺术。